ELSEVIER

Physica A 225 (1996) 157–163

PHYSICA A

The storage capacity of the complex phasor neural network

Zhenxiang Chen, Jianwei Shuai, Jincheng Zheng, Riutang Liu, Boxi Wu

Physics Department, Xiamen University, Xiamen 361005, China

Received 1 November 1995

Abstract

In this paper, the storage capacity of the Q-state complex phasor neural network is analysed with the signal-to-noise theory. The results indicate that the storage capacity of the model approaches that of the Hopfield model if the number Q is small; while the storage capacity is proportional to Q^{-2} if Q is large.

Keywords: Neural network; Complex phasor; Storage capacity

1. Introduction

One of the important applications of the neural computer of the future is to recognise the multistate patterns associationally. But until recently most of the theoretical attention was concentrated on the models having either binary variables or scalars constrained to the unit interval by sigmoidal nonlinearities [1–7]. Regardless of whether the interest lies in application or theory, it is rather important to set up the multistate neural network models to process the multistate gray or color patterns. Noest proposed that the neuron state could be a complex number, and set up the complex phasor neural networks [8,9]. For discrete state phasor neural networks, the Q-state signals are represented by Q discrete phasor variables uniformly distributed on the unit circle. Kanter suggested the Q-state Potts-glass neural network model [10], in which each neuron is viewed as a Pott spin and defined as a vector, pointing in Q directions which span a hypertetrahedron in R^Q-space. The Q-state integer neural network model [11] is discussed by Rieger, in which Q-state signals are represented by Q integers. By introducing the 2^n-element number into the Hopfield model, Shuai suggested the 2^n-element number neural network model [12,13].

SSDI 0378-4371(95)00408-4

The discrete Q-state complex neural network model [9,14] is an extended Hopfield model [1] and the four-state complex neural network model [15]. Such a model can describe natural or artificial networks of biological, chemical, or electronic limit-cycle oscillators with Q-fold instead of circular symmetry, or similar optical computing devices using a phase-encoded data representation. In any event, it is very likely even now that technical implementations of phasor models in optical hardware are feasible [9,14]. The complex phasor model is a more novel but also less studied model compared with other multistate models [10,11]. One of reasons is that there is no suitable treatment for the complex number with equilibrium statistical mechanics. But besides a statistical mechanics prescription [2-4], the signal-to-noise theory is also a popular treatment of the neural network [5-7]. In this paper, after a simple introduction of the complex phasor neural network model with the Dirac symbol representation, the signal-to-noise theory is used to analyse the storage capacity of the model.

2. The Dirac expression of the model

By introducing the complex plane rotation into the neural network model, the complex phasor model [9] can be set up. The state of the neuron is defined as a rotation operator and can be expressed as follows:

$$S = \mathrm{e}^{i(n\alpha+\beta)} \equiv [n]. \tag{1}$$

Here $\alpha = 2\pi/Q$, in which Q is the constancy integer, and the number $n = 0, 1, \dots, Q-1$. The angle β is the constancy angle, and $0 \le \beta < 2\pi/Q$. One can see that the values of the neuron are Q dots uniformly distributed in the unit circle, and simply expressed as $[n]$.

Suppose there are N neurons and M patterns $S^{\mu} = |\mu\rangle (\mu = 1, 2, \dots, M)$ stored in the network. Here with the representation of the Dirac symbol, the stored patterns are regarded as basic vectors. The synaptic connection matrix is given as the sum of the projection operators of basic vectors:

$$J = \sum_{\mu=1}^{M} |\mu\rangle\langle\mu|. \tag{2}$$

One can easily see that the connection matrix J is a Hermitian matrix: $J = J^{+}$. If a Q-state pattern $|S\rangle$ is put into the network, the energy of the model [8,14] is $E = \langle S|J|S\rangle$. The dynamics of the model depends only on the effective local field vector $|H\rangle$ expressed as

$$|H\rangle = J|S\rangle. \tag{3}$$

Actually the effective local field H_i is not the normalized rotation operator. The normalized local fields h_i can be simply expressed as $h_i = H_i/|H_i| \equiv [x_i]$. Here x_i is a real number, and the rotation angle of the local fields is expressed as $x_i\alpha+\beta$. The nonlinear dynamics of the neuron is defined as follows:

$$S_i(t+1)=\Theta[x_i]=[n_i], \qquad \text{if } -\tfrac{1}{2}\le n_i-x_i<\tfrac{1}{2}. \tag{4}$$

Obviously, we have $n_i=0,1,\dots,Q-1$. One can see that the function Θ transforms the local fields $h_i(t)$ to its nearest neuron state.

3. The storage capacity of the model

Now the storage capacity of the complex phasor neural network is discussed with the signal-to-noise theory with the thermodynamic limit in which $N,M\to\infty$ and the storage capacity $\gamma=M/N$ finite. If one of the stored patterns $|\nu\rangle$ is put into the network, the local field vector can be expanded as follows:

$$|H\rangle=N|\nu\rangle+\sum_{\mu\neq\nu}(\langle\mu|\nu\rangle)|\mu\rangle. \tag{5}$$

Clearly, in the expanded expresstion the local field is divided into two parts: the first term is the signal while the second term the noise. The noise term can be summed up as a complex number $K\exp(i\omega)$. Then the local fields can be rewritten in terms of each component with exponent representation respectively as

$$H_i=N\exp(i\theta_i^\nu)+K_i\exp(i\omega_i). \tag{6}$$

In the signal term, θ_i^ν is the rotation angle related with the ith pixel in the νth stored pattern and N is the weight of iteration to the state $|\nu\rangle$. Due to the random character of the stored patterns and their independence of each other, i.e, $\langle\nu|\mu\rangle\approx N\delta_{\nu\mu}$, it is reasonable to suppose that the noise term K_i is the module whose average value is 0 and the mean square error can be expressed as $\sigma=\sqrt{N(M-1)}\approx\sqrt{NM}$ containing the Gauss distribution as follows

$$f(r)=\sqrt{\frac{2}{\pi\sigma}}\exp\left(-\frac{r^2}{2\sigma^2}\right). \tag{7}$$

The angle ω_i is uniformly distributed in the angle range $[0,2\pi)$.

Because the model is symmetric under the rotation operator, without loss of generality, we can suppose that $\theta_i^\nu=0$, then

$$H_i=N+K_i\exp(i\omega_i)=|H_i|\exp(i\phi_i). \tag{8}$$

According to Eq. (4), the condition that the neuron i can be iterated correctly is $-\pi/Q\le\phi_i<\pi/Q$. Based on the view of geometry, which is shown in Fig. 1, N is expressed as the vector $\vec{\mathrm{OA}}$ on the x-axis. The noise term $K_i\exp(i\omega_i)$ is expressed as the vector $\vec{\mathrm{AB}}$. The length of the vector $\vec{\mathrm{AB}}$ follows a Gauss distribution, and the angle $\angle x\mathrm{AB}$ is ω_i. Then H_i is the sum of the vectors $\vec{\mathrm{OA}}$ and $\vec{\mathrm{AB}}$, i.e, $\vec{\mathrm{OB}}=\vec{\mathrm{OA}}+\vec{\mathrm{AB}}$. Now according to this geometry, the condition $-\pi/Q \leq \phi_i < \pi/Q$ means that $-\pi/Q \leq \angle\mathrm{AOB} < \pi/Q$, i.e, point B must fall into the sector $\angle$COD.

So the probability that the neuron i is able to iterate correctly equals the probability that point B falls into the sector $\angle$COD, and then is expressed as:

$$P = \int_0^R f(r)\,\mathrm{d}r + \int_R^N \left(1 - \frac{2}{\pi}\arccos(R/r)\right) f(r)\,\mathrm{d}r + \int_N^\infty \left(1 - \frac{1}{\pi}\arccos(R/N) - \frac{1}{\pi}\arccos(R/r)\right) f(r)\,\mathrm{d}r. \tag{9}$$

Here, the first term is expressed as the probability that point B falls in the circular area (1) with center at A and radius R, as shown in Fig. 2 shaded with dotted lines. The second term is expressed as the probability that the point B falls in the area (2) shown in Fig. 2 shaded with inverse-oblique lines, consisting of the intersection of the sector $\angle$COD and the circlular band with center at point A, inner radius R, and outer radius N. The third term is expressed as the probability that the point B falls into the area (3) shown in Fig. 2 shaded with oblique lines, consisting of the intersection of the sector $\angle$COD and the circlular band area with center at point A, inner radius N, and outer

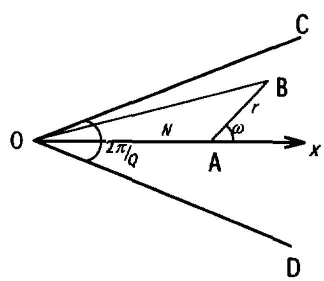

Fig. 1. The geometrical interpretation of the Eq. (8). The local field H_i is the vector $\vec{OB} = \vec{OA} + \vec{AB}$. To iterate correctly means that the spot B must fall into the sector $\angle$COD.

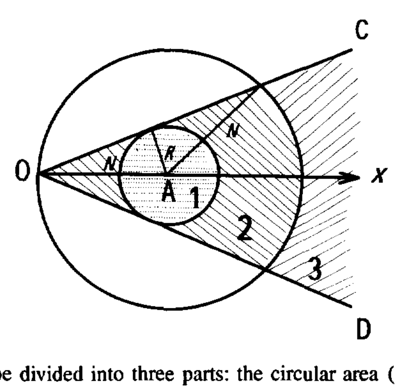

Fig. 2. The probability P can be divided into three parts: the circular area (1) shaded with dotted lines, the area (2) shaded with inverse-oblique lines, and the area (3) shaded with oblique lines.

radius ∞. Here, $R = N\sin(\pi/Q)$ is the radius of the circle tangent to sector $\angle$COD shown in Fig. 2. The probability P can be rewritten as the follows:

$$P = \int_0^\infty f(r)\,\mathrm{d}r - \int_R^N \frac{2}{\pi}\arccos(R/r)\,f(r)\,\mathrm{d}r$$
$$- \int_N^\infty \frac{1}{\pi}\arccos(R/N)\,f(r)\,\mathrm{d}r - \int_N^\infty \frac{1}{\pi}\arccos(R/r)\,f(r)\,\mathrm{d}r. \tag{10}$$

Especially when $Q = 2$, we can let the state of the neuron be $S = \pm 1$. If the boundary between two evolution areas is the i-axis in the dynamics equation, the neural network is just the Hopfield neural network model [1]. According the signal-to-noise theory [5,6], the storage capacity of the Hopfield model is $\alpha = \frac{1}{2}\ln N$.

Now the expression of the storage capacity of the complex phasor neural network with $Q > 2$ is discussed. If N is very large, we can write

$$\int_N^\infty \frac{1}{\pi}\arccos(R/N)\,f(r)\,\mathrm{d}r \approx \int_N^\infty \frac{1}{\pi}\arccos(R/r)\,f(r)\,\mathrm{d}r. \tag{11}$$

Then the probability of the network can be expressed as

$$P = 1 - \frac{2}{\pi}\int_R^\infty \arccos(R/r)\,f(r)\,\mathrm{d}r. \tag{12}$$

When r varies from R to ∞, $\arccos(R/r)$ varies from 0 to $\pi/2$; and the computer numerical simulation result indicates that the change of P mainly depends on the change of the function $f(r)$. So the function $\arccos(R/r)$ can be replaced by a constant coefficient A with $1/2 < A < \pi/2$. We can move A outside the definite integral. Therefore the expression of the probability P can be approximated by

$$P = 1 - \frac{2A}{\pi}\int_R^\infty f(r)\,\mathrm{d}r = 1 - \frac{A}{\pi}\left[1 - \sqrt{1 - \exp\left(-\frac{R^2}{2\sigma^2}\right)}\right]$$
$$= 1 - \frac{A}{\pi}\left\{1 - \sqrt{1 - \exp\left[\left(-\frac{1}{2\gamma}\right)\sin^2\frac{\pi}{Q}\right]}\right\}. \tag{13}$$

The error probability is $\rho = 1 - P$. The condition of the single neuron iterating correctly is $\rho \to 0$. If the number of error components follows approximately the Poisson distribution, then the probability of the correct components, i.e, the probability that $|S\rangle$ is indeed a stable attractor, is given approximately by the expression $\kappa = \exp(-N\rho)$. Now suppose we require that this probability be a fixed number very near 1; then inverting the preceding expression one can get $\rho = C/N$, where $C = -\ln\kappa$. This means that the storage capacity of the network can be obtained as:

$$\gamma = \frac{\sin^2(\pi/Q)}{2\ln(NA/\pi C)}. \tag{14}$$

Now as C, A are fixed, the relationship of the storage capacity with the neuron number and the neuron state value can be written as

$$\gamma \propto \frac{\sin^2(\pi/Q)}{2\ln N}. \tag{15}$$

The result indicates that the storage capacity reduces with the increasing of Q. We can interpret that result by considering the geometry shown in Fig. 1: as Q increases, the area of the sector $\angle$COD decreases, i.e, the probability that point B falls into the sector $\angle$COD decreases.

As we can see, the storage capacity approaches that of the Hopfield model, i.e, $\gamma \to \frac{1}{2}\ln N$ as $Q \to 2$. For practical purposes, the recognition the high precision patterns commonly used in computers is more important. For the complex neural network model with large neuron state value Q, the storage capacity is

$$\gamma \propto \frac{\pi^2}{2Q^2\ln N}. \tag{16}$$

The result indicates that the storage capacity is inversely proportional to Q^2 with Q large. An interpretation of this result is given as follows: When Q is large, the area of the sector $\angle$COD is very small. Therefore one can ignore the second and the third terms of Eq. (9), i.e,

$$P = \int_0^R f(r)\,\mathrm{d}r = \sqrt{1-\exp(-\pi^2/2\gamma Q^2)}. \tag{17}$$

To keep the probability $P \to 1$ leads to $\gamma \propto Q^{-2}$.

The results obtained here are quite different from those of Noest. By analysing the evolution of the overlaps, which is defined as the similarity of a network state to the patterns, Noest obtains in Ref. [9] the maximal memory capacity of the model as follows:

$$\gamma = (Q^2/4\pi)\sin^2(\pi/Q). \tag{18}$$

Then the maximal memory capacity converges for large Q to the value $\pi/4$.

Actually, from the different results obtained for the storage capacity of the Hopfield model using a different analysis, namely signal-to-noise theory [5,6] and statistical dynamics [2], one can see that the signal-to-noise theory defines a more restrictive approach to the memory capacity. It requires the stored patterns to be attractors of the network, provided no errors are introduced by interference among them.

4. Conclusion

In this paper, the storage capacity of the discrete complex phasor neural network model is discussed based on the signal-to-noise theory. A more restrictive result is obtained for the storage capacity. As we know, there exists no suitable treatment of complex numbers with equilibrium statistical mechanics providing a popular approach to the real neural network models [2–4]. We hope that a complex-number statistical approach can be set up, not only to analyse the complex neural network, but also to enrich the statistical mechanics, offering some new ideas or results for the study of real neural networks and spin glasses.

5. Acknowledgements

Funding for this work was provided in part by the National Natural Science Important Foundation under grant No. 19334032, and in part by National Natural Science High Technology Research Foundation under grant No. 69481003.

References

[1] J.J. Hopfield, Neural networks and physical systems with emergent collective computational abilities, Proc. Natl. Acad. Sci. USA 79 (1982) 2554.
[2] D.J. Amit, H. Gutfreund and H. Sompolinsky, Storing infinite numbers of patterns in a spin glass model of neural networks, Phys. Rev. Lett. 55 (1985) 1530.
[3] E. Scacciatelli and B. Tirozzi, Fluctuation of the free energy in the Hopfield model, J. Stat. Phys. 67 (1992) 981.
[4] A. Prugel-Bennett and J.L Shapiro, Statistical mechanics of unsupervised Hebbian learning, J. Phys. A 26 (1993) 2343.
[5] A.D. Bruce, E.J. Gardner and D.J. Wallace, Dynamics and statistical mechanics of the Hopfield model, J. Phys. A 20 (1987) 2909.
[6] R.J. Mceliece, E.D. Posner, E.R. Rodemich and S.S. Venkatesh, The capacity of the Hopfield associative memory, IEEE Tran. Inform. Theory IT 33 (1987) 461.
[7] Shiino and T Fukai, Self-consistent signal-to-noise analysis and its application to analogue neural networks with asymmetric connections, J. Phys. A 25 (1992) L375.
[8] A.J. Noest, Associative memory in sparse phasor neural network, Europhys. Lett. 6 (1988) 469.
[9] A.J. Noest, Discrete-state phasor neural networks, Phys. Rev. A 38 (1988) 2196.
[10] I. Kanter, Potts-glass models of neural networks, Phy. Rev. A 37 (1988) 2739.
[11] H. Rieger, Storing an extensive number of grey-toned patterns in a neural network using multistate neurons, J. Phys. A 23 (1990) L1273.
[12] J.W. Shuai, Z.X. Chen, R.T. Liu and B.X. Wu, The 2^n-element number neural network model: recognition of the multistate patterns, J. Mod. Opt. 42 (1995) 1179.
[13] J.W. Shuai, Z.X. Chen, R.T. Liu and B.X. Wu, The Hamilton neural network model, Physica A 216 (1995) 20.
[14] G.Y. Sirat, A.D. Maruani and R.C. Chevallier, Gray level neural networks, Appl. Opt. 28 (1989) 414.
[15] C.H. Zhou and L.R. Liu, The complex Hopfield model, Opt. Commun. 103 (1993) 29.